

p#17



ENTERED

1600

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/455,978BDATE: 03/27/2002  
TIME: 14:52:17Input Set : A:\H1020011.app  
Output Set: N:\CRF3\03272002\I455978B.rawRECEIVED  
APR 02 2002  
TECH CENTER 1600/2800

3 <110> APPLICANT: Alam, Maqsudul  
4 Larsen, Randy  
6 <120> TITLE OF INVENTION: HEME PROTEINS HEMAT-HS AND HEMAT-BS AND THEIR USE IN  
7 MEDICINE AND MICROSENSORS  
9 <130> FILE REFERENCE: 201040/1020  
11 <140> CURRENT APPLICATION NUMBER: 09/455,978B  
12 <141> CURRENT FILING DATE: 1999-12-06  
14 <160> NUMBER OF SEQ ID NOS: 86  
16 <170> SOFTWARE: PatentIn Ver. 2.1  
18 <210> SEQ ID NO: 1  
19 <211> LENGTH: 1470  
20 <212> TYPE: DNA  
21 <213> ORGANISM: Halobacterium salinarum  
23 <400> SEQUENCE: 1  
24 atgagcaacg ataatgacac tctcgtgacc gccgacgttc ggaacgggat cgacgggcac 60  
25 gcactcgcgg acccgatcgg cctcgacgag gccgagatcg cgtggcggt gtcgttcacc 120  
26 gggatcgacg acgacacgat ggccgcgctc gccgccgaac agccgctgtt cgaagccacc 180  
27 gcggacgcgc tggtagccga cttctacgac cacttgaggt cctacgagcg cacacaggac 240  
28 ctcttcgcga actccacgaa gaccgtcgag caactcaaag agacgcaggc cgagtacttg 300  
29 ctgggcctcg ggcgcggcga gtaacacacc gagtacgcg cccagcgcgc ccgtatcggg 360  
30 aagatacacg acgtgctcgg gctcggaccg gacgtctatc tgggcgcgta cacgcgatac 420  
31 tacacggggc tgttgagcgc gcttgccgac gacgtggctg ccgaccgcgg cgaggaggcg 480  
32 gccgcgcgcg tcgacgaact cgtggcccgg ttcttgccga tgttgaagct gttgaccttc 540  
33 gatcagcaga tcgcaatgga cacctacatc gactcgtacg cccagcgcct ccacgcagag 600  
34 atcgacagcc gccaggagtt ggccaacgcg gtcgccacgc acgtggaagc accgctgtcc 660  
35 tcgctggagg cgacctcgca ggacgtcgcc gagcgcaagg acacgatgcg ggcccgcacc 720  
36 gacgaccagg tcgaccgcat ggctgacgtc agccgtgaga tatccagcgt gtccgcgagc 780  
37 gtcgaggagg tcgcctcgac ggccgacgac gtccgcggga ccagcgagga cgccgaggcg 840  
38 ctggcccagc agggcgaggc ggccgcgcgac gacgcgctcg ccacgatgac cgacatcgac 900  
39 gaggcgaccg acggcgctac cgcgggcgct gaacagctcg gcgagcgcgc cgccgacgtc 960  
40 gaatcagtga ccggcgtgat cgacgacatc gccgagcaga cgaacatgct ggcgctgaac 1020  
41 gcgtccatcg aggcgcgccg cgcgggggag gcgggcgagg ggtttgcggt cgtcgccgac 1080  
42 gaggtcaagg cctcgcgca ggagtcgcgc gagcagtcga cgcgcgctga ggagctcgtc 1140  
43 gagcagatgc aggcggagac cgaggagacg gtcgaccagt tggacgaggt caaccagcgc 1200  
44 atcggcgagg gcgtcgagcg cgtcgaggag gcgatggaga cctccagga gatcaccgac 1260  
45 gccgtcgagg acgcgcgaag cgggatgcag gaggtgtcga cggcgaccga cgaacaggcg 1320  
46 gtgagcaccg aggaggtcgc cgagatggtc gacggtgtcg acgaccgcgc gggcgagatc 1380  
47 gcggccgccc tcgatgacat cgcgcgacgc accgatcagc aggtccggac cgtcgaggag 1440  
48 gtccgcgaga cggtcggcaa gctcagctag 1470  
51 <210> SEQ ID NO: 2  
52 <211> LENGTH: 489  
53 <212> TYPE: PRT  
54 <213> ORGANISM: Halobacterium salinarum

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/455,978B

TIME: 14:52:17

Input Set : A:\H1020011.app

Output Set: N:\CRF3\03272002\I455978B.raw

56 &lt;400&gt; SEQUENCE: 2

```

57 Met Ser Asn Asp Asn Asp Thr Leu Val Thr Ala Asp Val Arg Asn Gly
58   1           5           10           15
60 Ile Asp Gly His Ala Leu Ala Asp Arg Ile Gly Leu Asp Glu Ala Glu
61           20           25           30
63 Ile Ala Trp Arg Leu Ser Phe Thr Gly Ile Asp Asp Asp Thr Met Ala
64           35           40           45
66 Ala Leu Ala Ala Glu Gln Pro Leu Phe Glu Ala Thr Ala Asp Ala Leu
67           50           55           60
69 Val Thr Asp Phe Tyr Asp His Leu Glu Ser Tyr Glu Arg Thr Gln Asp
70   65           70           75           80
72 Leu Phe Ala Asn Ser Thr Lys Thr Val Glu Gln Leu Lys Glu Thr Gln
73           85           90           95
75 Ala Glu Tyr Leu Leu Gly Leu Gly Arg Gly Glu Tyr Asp Thr Glu Tyr
76           100          105          110
78 Ala Ala Gln Arg Ala Arg Ile Gly Lys Ile His Asp Val Leu Gly Leu
79           115          120          125
81 Gly Pro Asp Val Tyr Leu Gly Ala Tyr Thr Arg Tyr Tyr Thr Gly Leu
82           130          135          140
84 Leu Asp Ala Leu Ala Asp Asp Val Val Ala Asp Arg Gly Glu Glu Ala
85  145          150          155          160
87 Ala Ala Ala Val Asp Glu Leu Val Ala Arg Phe Leu Pro Met Leu Lys
88           165          170          175
90 Leu Leu Thr Phe Asp Gln Gln Ile Ala Met Asp Thr Tyr Ile Asp Ser
91           180          185          190
93 Tyr Ala Gln Arg Leu His Asp Glu Ile Asp Ser Arg Gln Glu Leu Ala
94           195          200          205
96 Asn Ala Val Ala Thr His Val Glu Ala Pro Leu Ser Ser Leu Glu Ala
97           210          215          220
99 Thr Ser Gln Asp Val Ala Glu Arg Thr Asp Thr Met Arg Ala Arg Thr
100  225          230          235          240
102 Asp Asp Gln Val Asp Arg Met Ala Asp Val Ser Arg Glu Ile Ser Ser
103           245          250          255
105 Val Ser Ala Ser Val Glu Glu Val Ala Ser Thr Ala Asp Asp Val Arg
106           260          265          270
108 Arg Thr Ser Glu Asp Ala Glu Ala Leu Ala Gln Gln Gly Glu Ala Ala
109           275          280          285
111 Ala Asp Asp Ala Leu Ala Thr Met Thr Asp Ile Asp Glu Ala Thr Asp
112           290          295          300
114 Gly Val Thr Ala Gly Val Glu Gln Leu Gly Glu Arg Ala Ala Asp Val
115  305          310          315          320
117 Glu Ser Val Thr Gly Val Ile Asp Asp Ile Ala Glu Gln Thr Asn Met
118           325          330          335
120 Leu Ala Leu Asn Ala Ser Ile Glu Ala Ala Arg Ala Gly Glu Ala Gly
121           340          345          350
123 Glu Gly Phe Ala Val Val Ala Asp Glu Val Lys Ala Leu Ala Glu Glu
124           355          360          365
126 Ser Arg Glu Gln Ser Thr Arg Val Glu Glu Leu Val Glu Gln Met Gln
127           370          375          380

```

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/455,978B

TIME: 14:52:17

Input Set : A:\H1020011.app

Output Set: N:\CRF3\03272002\I455978B.raw

129 Ala Glu Thr Glu Glu Thr Val Asp Gln Leu Asp Glu Val Asn Gln Arg  
 130 385 390 395 400  
 132 Ile Gly Glu Gly Val Glu Arg Val Glu Glu Ala Met Glu Thr Leu Gln  
 133 405 410 415  
 135 Glu Ile Thr Asp Ala Val Glu Asp Ala Ala Ser Gly Met Gln Glu Val  
 136 420 425 430  
 138 Ser Thr Ala Thr Asp Glu Gln Ala Val Ser Thr Glu Glu Val Ala Glu  
 139 435 440 445  
 141 Met Val Asp Gly Val Asp Asp Arg Ala Gly Glu Ile Ala Ala Ala Leu  
 142 450 455 460  
 144 Asp Asp Ile Ala Asp Ala Thr Asp Gln Gln Val Arg Thr Val Glu Glu  
 145 465 470 475 480  
 147 Val Arg Glu Thr Val Gly Lys Leu Ser  
 148 485

151 &lt;210&gt; SEQ ID NO: 3

152 &lt;211&gt; LENGTH: 1390

153 &lt;212&gt; TYPE: DNA

154 &lt;213&gt; ORGANISM: Bacillus subtilis

156 &lt;400&gt; SEQUENCE: 3

157 atgttatttta aaaaagacag aaaacaagaa acagcttact tttcagattc aaacggacaa 60  
 158 caaaaaaacc gcattcagct cacaacaaca catgcagatg tcaaaaaaca gctcaaaatg 120  
 159 gtcagggttg gagatgctga gctttatgtg ttagagcagc ttcagccact cattcaagaa 180  
 160 aatatcgtaa atatcgctga tgcgttttat aaaaaccttg accatgaaag ctcattgatg 240  
 161 gatatacatta atgatcacag ctcagttgac cgcttaaaac aaacggttaa acggcatatt 300  
 162 caggaaatgt ttgcaggcgt tatcgatgat gaatttattg aaaagcgtaa ccgaatcgcc 360  
 163 tccatccatt taagaatcgg ccttttgcca aaatgggtata tgggtgcgtt tcaagagctc 420  
 164 cttttgtcaa tgattgacat ttatgaagcg tccattacaa atcagcaaga actgctaaaa 480  
 165 gccattaaag caacaacaaa aatcttgaac ttagaacagc agcttgtcct tgaagcgttt 540  
 166 caaagcgagt acaaccagac ccgtgatgaa caagaagaaa agaaaaacct tcttcatcag 600  
 167 aaaattcaag aaacctctgg atcgattgcc attctgtttt cagaaacaag cagatcagtt 660  
 168 caagagcttg tggacaaatc tgaaggcatt tctcaagcat ccaaagccgg cactgtaaca 720  
 169 tccagcaactg ttgaagaaaa gtcgatcggc ggaaaaaaag agctagaagt ccagcaaaaa 780  
 170 cagatgaaca aaattgacac aagccttgtc caaatcgaaa aagaaatggg caagctggat 840  
 171 gaaatcgcgc agcaaattga aaaaatcttc ggcacgtcga caggcatagc tgaacaaaca 900  
 172 aaccttctct cgctcaatgc atctattgaa tccgcccgcg ccggagaaca cggcaaaggc 960  
 173 tttgctgtcg tggcaaatga agtgcggaag ctttctgagg atacgaaaaa aaccgtctct 1020  
 174 actgtttctg agcttgtgaa caatacgaat acacaaatca acattgtatc caagcatatc 1080  
 175 aaagacgtga atgagctagt cagcgaaagt aaagaaaaaa tgacgcaaat taaccgctta 1140  
 176 ttcgatgaaa tcgtccacag catgaaaatc agcaaagagc aatcaggcaa aatcgacgtc 1200  
 177 gatctgcaag cttttcttgg agggcttcag gaagtcagcc gcgcccgttt ccatgtggcc 1260  
 178 gcttccgctt attcgcttgc catcctgaca gaagaataac catcaaaaac cggctctgcca 1320  
 179 tacggccggt ttttttgcgt tcattatgta aacttaaatt aaaaatcagt tgacataata 1380  
 180 attacctgca 1390

183 &lt;210&gt; SEQ ID NO: 4

184 &lt;211&gt; LENGTH: 432

185 &lt;212&gt; TYPE: PRT

186 &lt;213&gt; ORGANISM: Bacillus subtilis

188 &lt;400&gt; SEQUENCE: 4

189 Met Leu Phe Lys Lys Asp Arg Lys Gln Glu Thr Ala Tyr Phe Ser Asp

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/455,978B

TIME: 14:52:17

Input Set : A:\H1020011.app

Output Set: N:\CRF3\03272002\I455978B.raw

```

190      1              5              10              15
192 Ser Asn Gly Gln Gln Lys Asn Arg Ile Gln Leu Thr Asn Lys His Ala
193              20              25              30
195 Asp Val Lys Lys Gln Leu Lys Met Val Arg Leu Gly Asp Ala Glu Leu
196              35              40              45
198 Tyr Val Leu Glu Gln Leu Gln Pro Leu Ile Gln Glu Asn Ile Val Asn
199              50              55              60
201 Ile Val Asp Ala Phe Tyr Lys Asn Leu Asp His Glu Ser Ser Leu Met
202 65              70              75              80
204 Asp Ile Ile Asn Asp His Ser Ser Val Asp Arg Leu Lys Gln Thr Leu
205              85              90              95
207 Lys Arg His Ile Gln Glu Met Phe Ala Gly Val Ile Asp Asp Glu Phe
208              100              105              110
210 Ile Glu Lys Arg Asn Arg Ile Ala Ser Ile His Leu Arg Ile Gly Leu
211              115              120              125
213 Leu Pro Lys Trp Tyr Met Gly Ala Phe Gln Glu Leu Leu Ser Met
214 130              135              140
216 Ile Asp Ile Tyr Glu Ala Ser Ile Thr Asn Gln Glu Leu Leu Lys
217 145              150              155              160
219 Ala Ile Lys Ala Thr Thr Lys Ile Leu Asn Leu Glu Gln Gln Leu Val
220              165              170              175
222 Leu Glu Ala Phe Gln Ser Glu Tyr Asn Gln Thr Arg Asp Glu Gln Glu
223              180              185              190
225 Glu Lys Lys Asn Leu Leu His Gln Lys Ile Gln Glu Thr Ser Gly Ser
226              195              200              205
228 Ile Ala Asn Leu Phe Ser Glu Thr Ser Arg Ser Val Gln Glu Leu Val
229 210              215              220
231 Asp Lys Ser Glu Gly Ile Ser Gln Ala Ser Lys Ala Gly Thr Val Thr
232 225              230              235              240
234 Ser Ser Thr Val Glu Glu Lys Ser Ile Gly Gly Lys Lys Glu Leu Glu
235              245              250              255
237 Val Gln Gln Lys Gln Met Asn Lys Ile Asp Thr Ser Leu Val Gln Ile
238              260              265              270
240 Glu Lys Glu Met Val Lys Leu Asp Glu Ile Ala Gln Gln Ile Glu Lys
241              275              280              285
243 Ile Phe Gly Ile Val Thr Gly Ile Ala Glu Gln Thr Asn Leu Leu Ser
244 290              295              300
246 Leu Asn Ala Ser Ile Glu Ser Ala Arg Ala Gly Glu His Gly Lys Gly
247 305              310              315              320
249 Phe Ala Val Val Ala Asn Glu Val Arg Lys Leu Ser Glu Asp Thr Lys
250              325              330              335
252 Lys Thr Val Ser Thr Val Ser Glu Leu Val Asn Asn Thr Asn Thr Gln
253              340              345              350
255 Ile Asn Ile Val Ser Lys His Ile Lys Asp Val Asn Glu Leu Val Ser
256              355              360              365
258 Glu Ser Lys Glu Lys Met Thr Gln Ile Asn Arg Leu Phe Asp Glu Ile
259 370              375              380
261 Val His Ser Met Lys Ile Ser Lys Glu Gln Ser Gly Lys Ile Asp Val
262 385              390              395              400

```

## RAW SEQUENCE LISTING

DATE: 03/27/2002

PATENT APPLICATION: US/09/455,978B

TIME: 14:52:17


Input Set : A:\H1020011.app

Output Set: N:\CRF3\03272002\I455978B.raw

```

264 Asp Leu Gln Ala Phe Leu Gly Gly Leu Gln Glu Val Ser Arg Ala Val
265          405          410          415
267 Ser His Val Ala Ala Ser Val Asp Ser Leu Val Ile Leu Thr Glu Glu
268          420          425          430
274 <210> SEQ ID NO: 5
275 <211> LENGTH: 57
276 <212> TYPE: PRT
277 <213> ORGANISM: Artificial Sequence
279 <220> FEATURE:
280 <223> OTHER INFORMATION: Description of Artificial Sequence: Template
281     sequence
283 <220> FEATURE:
284 <221> NAME/KEY: UNSURE
285 <222> LOCATION: (4)..(57)
286 <223> OTHER INFORMATION: Xaa at positions 4, 10, 14, 15, 27, and 41-57 is
287     unknown
289 <400> SEQUENCE: 5
W--> 290 Ile Ile Lys Xaa Thr Val Pro Val Leu Xaa Glu His Gly Xaa Xaa Ile
291     1         5         10         15
W--> 293 Gly Gln Asp Val Leu Val Val Leu Ile Lys Xaa Asn Pro Glu Ile Gln
294     20         25         30
W--> 296 Glu Lys Phe Phe Phe Phe Lys His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
297     35         40         45
W--> 299 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
300     50         55
303 <210> SEQ ID NO: 6
304 <211> LENGTH: 55
305 <212> TYPE: PRT
306 <213> ORGANISM: Erwinia chrysanthemi
308 <400> SEQUENCE: 6
309 Ile Lys Ser Thr Ile Pro Leu Leu Ala Glu Thr Gly Pro Ala Leu Thr
310     1         5         10         15
312 Ala His Phe Tyr Gln Arg Met Phe His Asn Pro Glu Leu Lys Asp
313     20         25         30
315 Ile Phe Asn Met Ser Asn Gln Arg Asn Gly Asp Gln Arg Glu Ala Leu
316     35         40         45
318 Phe Asn Ala Ile Cys Ala Tyr
319     50         55
322 <210> SEQ ID NO: 7
323 <211> LENGTH: 56
324 <212> TYPE: PRT
325 <213> ORGANISM: Vitreoscilla stercoraria
327 <400> SEQUENCE: 7
328 Ile Ile Lys Ala Thr Val Pro Val Leu Lys Glu His Gly Val Thr Ile
329     1         5         10         15
331 Thr Thr Thr Phe Tyr Lys Asn Leu Phe Ala Lys His Pro Glu Val Arg
332     20         25         30
334 Pro Leu Phe Asp Met Gly Arg Gln Glu Ser Leu Glu Gln Pro Lys Ala
335     35         40         45

```


 Use of n and / or Xaa has been detected in the  
 Sequence Listing. Review the Sequence Listing  
 to ensure a corresponding explanation is present  
 in the <220> to <223> fields of each sequence  
 using n or Xaa.

## VERIFICATION SUMMARY

DATE: 03/27/2002

PATENT APPLICATION: US/09/455,978B

TIME: 14:52:18

Input Set : A:\H1020011.app

Output Set: N:\CRF3\03272002\I455978B.raw

L:290 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:293 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:296 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:299 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:753 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30  
L:756 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30  
L:935 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41  
L:953 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42  
L:1062 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50  
L:1701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82  
L:1704 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82  
L:1723 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83  
L:1745 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84  
L:1764 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85  
L:1786 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86